Osmat A. Jefferson
Director of Lens Product Development, Cambia.
Professor of Science and Technology,
Queensland University of Technology (QUT)
Australia
CURRENT OFFERINGS

Discovery, Analytics, And Management Tools

**Collections** (dynamic): Based on Scholarly Works and Patents Lens searches and analyses.

**API & Data Facility**: Scholarly API, PatSeq bulk data, Patent API*

**In4M**: International Industry & Innovation Influence Mapping

**Reports**: Assemble your saved queries and collections with other knowledge in a dynamic and interactive report*

* coming soon
Lens Labs is an open collaboration that exposes innovation datasets and engages academic communities to build a more transparent innovation ecology.

https://www.lens.org/lens/labs
COVID-19 Datasets
Free and open datasets of patent documents, scholarly research works metadata and biological sequences from patents.

Bulk Patent Data
US patent bulk data is available on request for non-commercial use under a CC-BY-NC licence.

Lens Patent API
Patent API including access to abstracts, claims, owners and legal status (coming soon)

Lens Scholarly API
Access the Lens Scholarly API, including bibliographic metadata, cited and citing scholarly works and citing patents.

Patseq Data
Access to the human genome map of patent sequences for academic use, and all 350M patent sequences disclosed in the patent literature.

GitHub Repositories
Major Research Contributions from HHMI Authors.

Massachusetts Institute of Technology
MIT is a private land-grant, sea-grant, and space-grant university in Cambridge, Massachusetts, United States.

Downloads:
- Scholarly Works [372 MB]
- Works cited by patents [79 MB]
- Citing Patents [3.9 GB]
- Patents [300 MB]
Community-based features

User surveys, feedback, and discussions with diverse institutions have contributed to several Lens features.

Here is the development pipeline planned for 2020-2021

Your role and participation matter!
October 2020

Lens Patent Data

126.4 Million Patent Records:

- 105 jurisdictions
- 69.4M patent families
- 727k biological patents
- 366.5M patent sequences
A patent is best represented by a dynamic meta record

Besides conferring a limited exclusive right on an invention, a patent is an accumulated body of scientific, technical, legal, business and industrial knowledge that is continually evolving based on other related knowledge artefacts.

To gain insight on an invention in an innovation pathway, it is critical to access, analyse, link and monitor its varied contextual metadata; the patent file wrapper, its family members, the citations, the office actions, the legal challenges, and other relevant data elements.
THE LENS

Meta Record Strategy

**Patent Meta Record Overview**

1. Patent document records from Patent Transfer Offices (PTOs) and the European Patent Office’s patent bibliographic database (DOCDB) are ingested separately into the system. Each record is assigned a unique Lens ID.

2. The ingested records are matched using their publication identifiers.

3. The individual record data is merged to create a Patent Metadata Record (PMR).

4. Supplementary data sets such as legal events, ownership transfers and patent family. Lens SMR and Lens PaperIdentifiers are added to the PMR.

5. The Lens ID of the first cluster member processed is chosen to be the PMR identifier in the Lens system.

**Scholarly Meta Record Overview**

1. Citations strings are resolved to external records such as PubMed, CrossRef and MAG. Each of these records are assigned a unique Lens ID.

2. The records are clustered by matching or linking identifiers and by metadata similarity rules.

3. The metadata for the records in each cluster is merged to create a Scholarly Metadata Record (SMR).

4. Supplementary data sets such as ORCID, CORE, DOAJ and Unpaywall are matched and added to the SMR.

5. A representative Lens ID of a cluster member is chosen to be the single SMR identifier in the Lens system. However, all individual cluster Lens IDs resolve to their associated SMR, allowing SMR cluster structure to change over time.

https://osf.io/preprints/lissa/t56yh/
Patent MetaRecord (PMR) Store and Index
### API & Data

**Lens Patent API**

The Lens Patent API is now in Beta testing phase and undergoing development. Email support@lens.org for more information.

Alternatively, Take the API Survey to improve the API.

**Request Notification for Patent API**

<table>
<thead>
<tr>
<th>Classification Field</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>classifications_cpc.action_date</td>
<td>Date</td>
<td>Issue date of the patent document. yyyy-MM-dd</td>
</tr>
<tr>
<td>classifications_cpc.classification_data_source</td>
<td>String</td>
<td>H - Human, M - Machine, G - Generated, C - Classified</td>
</tr>
<tr>
<td>classifications_cpc.classification_status</td>
<td>String</td>
<td>B - Original, R - Reclassified</td>
</tr>
<tr>
<td>classifications_cpc.classification_symbol_position</td>
<td>String</td>
<td>F - First, L - Later</td>
</tr>
<tr>
<td>classifications_cpc.classification_value</td>
<td>String</td>
<td>I - Invention, A - Additional</td>
</tr>
<tr>
<td>classifications_cpc.first_symbol</td>
<td>String</td>
<td></td>
</tr>
<tr>
<td>classifications_cpc.generating_office</td>
<td>Country</td>
<td></td>
</tr>
<tr>
<td>classifications_cpc.later_symbol</td>
<td>String</td>
<td></td>
</tr>
<tr>
<td>classifications_cpc.sequence</td>
<td>Integer</td>
<td></td>
</tr>
<tr>
<td>classifications_cpc.source</td>
<td>Source</td>
<td>e.g. DOCDB</td>
</tr>
<tr>
<td>classifications_cpc.symbol</td>
<td>String</td>
<td>e.g. H04L47/2416</td>
</tr>
<tr>
<td>classifications_cpc.version_indicator</td>
<td>Date</td>
<td>e.g. 2006-01-01</td>
</tr>
</tbody>
</table>
PMR integrated and released to Production

- **PMR Store & Index:** Integrated PMR data model in a MongoDB store with Elastic search index including: new calculated fields, additional full text and Non patent literature citations, backward and forward patent citations, improved and consistent patent key logic, classification hierarchy, and weekly delta updates.

- **Legal Events:** An INPADOC legal events store and new calculated legal event fields including: Prosecution stage, legal status, owner history, legal events timeline.

- **Refactored Environment:** Refactored User Interface environment.

- **Patent API:** Patent API Integrated with the existing subscription model.
PMR 2021- Roadmap

- Continue INPADOC integration and inclusion of legal events
THE LENS

PMR 2021- Roadmap

- Patent family data store (family centric search)
- Expanded jurisdiction (Germany and Canada)
- Expanded NPLs (integration from the full text)
- Clustered Classification codes
The Lens Cooperative


Initially supported by the Bill & Melinda Gates Foundation, AP Sloan Foundation, The Rockefeller Foundation and Amazon Web Services.

The Lens Cooperative allows government agencies, research institutions, foundations and enterprise to join, support, prioritize, influence, use and benefit from a sustainable open infrastructure.

Launching in January 2021.
Beta members joining now.
Why is CRISPR-Cas9 a revolutionary technology?

It is fast and easy to implement, cheap and components are readily accessible. This versatility means that the technology can deliver a timely end product and can be used by many stakeholders.

https://link.lens.org/edYPNxark3b
SARS-CoV-2: Navigating information overload with Lens.org

Evidence must be used to guide interventions for the COVID-19 pandemic, but that evidence must be comprehensive, credible and shared for it to be effective. This report focuses on how global scientific research and patenting activity relevant to SARS-CoV-2, including its genetics and pathogenesis, can be discovered and made transparent, open, shareable and navigable to help inform how it could be translated into intervention options.

We demonstrate how the platform Lens.org can be used to conduct reliable systematic scholarly and patent searches, refine and build domain-based collections to share publicly, and set alerts to receive notifications for updates. We also provide some analyses/comments on some of these collections. Our aim is to enable researchers to find clarity into the discovery, analysis, and translation of COVID-19 scientific and patent knowledge into products, practices or services to help mitigate the current crisis.

Credit for background picture: a screenshot for SARS-CoV-2 from a photo in Immune Matter, im Spring 2020

Authored by: Osmat Azzam Jefferson, Simon Lang, Kenny Williams, Aaron Baillagh, Deniz Koehler, Ben Warren, Bernard Schellberg, Roshan Sharma, Richard Jefferson